

Beginning at page 6, line 3:

A3 The present invention is applied to an HA (hearing aid) 1 configured as shown for example in Figs. 1 and 2. This HA 1 is of a portable type and includes an HMD (head-mounted display) 2, and a computer unit 3 for speech recognition and for generation of the speech information etc., interconnected over an optical fiber cable 4, as shown in Fig. 1. The computer unit 3 is annexed to a support 5 worn on e.g., a user's waist and is driven by power supplied from a battery 6 annexed to the support 5, while driving the HMD 2.

Beginning at page 6, line 15:

A4 The display unit 7, arranged just ahead of the user's eye, demonstrates the meaning or the contents of the speech detected by the microphone for a user 8 and/or by the microphone for outside 11 which will be explained subsequently. It is also possible for the display unit 7 to demonstrate not only the meaning and the contents of the speech as described above but also the other information responsive to commands from the computer unit 3.

Beginning at page 7, line 7:

A5 The microphone for a user 8 and the microphone for outside 11 may be designed as processor for acoustic signals, such as acoustic echo-canceller, a microphone array or a variety of microphones, irrespective of the mounting positions depending on the user's operation. These microphones may be exemplified by these types (pressure-type, pressure-gradient-type, parametric-type, laser Doppler-type, bone-conduction-type, ultra-small two-way unit picking up air-conduction and bone-conduction sound (manufactured by NTT), non-directional-type, uni-directional (ultra-directional)-type, bi-directional-type, dynamic-type, condenser (electret)-type, zoom-type, stereo-type, MS stereo-type, wireless-type, ceramic-type and magnetic-type).

Beginning at page 8, line 21:

A6 The computer unit 3 starts the program stored in the recording medium, based on the electrical signals generated from the speech detected by the microphone for a user 8 and/or the microphone for outside 11 and performs speech recognition processing in the CPU to produce the results of recognition. This permits the computer unit 3 to obtain the

contents of the speech detected by the microphone for a user 8 and/or the microphone for outside 11 in the CPU.

Beginning at page 14, line 3:

A7

The speech information generating unit 23 works on, transforms or synthesizes etc. the results of recognition from the signal processor 22, depending on the bodily state, the using state or the using objectives of the user. The speech information generating unit 23 also performs the processing of presenting the speech detected by the microphone 21 to the user on the results of recognition or the results of recognition obtained on working etc.

Paragraph beginning at page 14, line 9:

A8

It is also possible for the speech information generating unit 23 to modify the speech information generated from the results of recognition to generate the new speech information. At this time, the speech information generating unit 23 annexes words or phrases more familiar to the user, based on the bodily state, the using state or the using objectives of the user, to improve further the speech recognition by the user. For example, if "BIG MAC" (Trademark) is inputted to the microphone 21, the speech information generating unit 23 doing this processing generates e.g., the speech information indicating "MACDONALD'S BIG MAC" (Trademark).

A9

Beginning at page 20, line 3:

The speech enhancement unit 25, adapted for outputting the speech to the user's ear, may be of the dynamic speech enhancement transducing system or electro-static speech enhancement transducing system (capacitor or electrostatic type), and of the shape of a headphone such as the open-air, closed, or in-the-ear type such as canal type. On the other hand, the speech enhancement unit 25 may be that used in a conventional HA, speech enhancement unit or sound collector speaker, may be of the type employing a magnetic loop or may be of a microphone speaker system employing fingers (prototype: WHISPER (NTT Docomo)). After all, the speech enhancement unit 25 outputting the speech from the user to the speaker may be any suitable conventionally used speech enhancement unit or speaker.

Beginning at page 29, line 5:

A 10 When synthesizing speech data, the speech information generating unit 23 is able to execute speech synthesis by rule irrespective of the speech contents, speech synthesis employing variable length unit to synthesize the smooth speech, prosody control to produce natural speech on synthesis or quantizing conversion for furnishing the speech with personality in order to generate the speech information (see "A perspective in automatic telephone interpretation" Advanced Telecommunication Res. Ins. ATR Advanced Tech. Series, Ohmsha, 177-209, 1994).

Beginning at page 30, line 18:

A 11 It is also possible for the speech information generating unit 23 not only to perform the processing of interpreting the language from the speech, using the result of recognition, and of constructing the speech information from the speech data using the so-interpreted language, but also to perform the processing of working on or transforming the language from other processing interpreted on the basis of the results of recognition. That is, it is possible for the present speech information generating unit 23 to construct the speech information and to perform the voice speed conversion of changing the rate of outputting the speech information to the speech enhance unit 25, such as by elongating the voiced domain by splitting or elongating the pitch domain, not working on the invoiced domain, or shortening the non-speech domain. That is, this voice speed conversion is performed by selecting appropriate voice speed depending on the user's state.

Beginning at page 34, line 10:

A 12 If the speech information generating unit 23 has heard an alarm bell, a siren of an ambulance car or a siren informing of oncoming tsunami, it not only demonstrates that effect but outputs from the speech enhancement unit 25 such announcement as "it's fire", "ambulance car has arrived" or "tsunami is oncoming" with a large volume, while demonstrating an image representing the fire, ambulance car or tsunami on the display unit 26.

Beginning at page 38, line 13:

A 13 It is moreover possible for the signal processor 22 and the speech information generating unit 23 to perform speech recognition only for a specified noise, in addition to generating the result of recognition only of the speech of the speaker to generate the speech information for presentation to the speech enhancement unit 25 and/or the display unit 26 and thence to the user. In sum, the signal processor 22 and the speech information generating unit 23 performs speech recognition processing on the input sound to transform the result of recognition depending on the bodily state, the using state or the using objectives of the user, to generate the speech information in expressions readily comprehensible for the user to output the generated speech information.

Beginning at page 47, line 6:

A 14 The HA 1 is not limited to an embodiment in which the speech recognition processing is executed by the signal processor 22 using the speech from the microphone 21 etc. For example, the speech recognition processing may also be executed using signals from a nasal sound sensor, respiratory air-stream sensor, neck-vibration sensor, bone vibrator, such as mouthpiece-type vibrator, worn by the user or by a person other than the user, or using signals from the microphone 21. Thus, with the HA 1, the rate of recognition by the signal processor 22 may be improved further by employing respective sensors in addition to the microphone 21.

Beginning at page 47, line 14:

A 15 The HA 1 may also be provided with a camera mechanism 29 for photographing a moving picture or a still image by a digital camera carrying e.g., an automatic focussing mechanism or a zooming function, to display it on the display unit 26, as shown in Fig.2. This camera mechanism 29 may be loaded as one with the display unit 7 of Fig.1. The camera mechanism 29 may use a digital camera.

Beginning at page 48, line 3:

A 16 The HA 1 demonstrates a photographed image on the display unit 26 from the camera mechanism 29 through a signal processing circuit. With the HA 1, the image which has photographed the speaker by the camera mechanism 29 is presented to the user to improve the user's recognition. It is possible with this HA 1 to output the

A 16
cont

photographed image through the communication circuit 27 to an external network. It is also possible with the HA 1 to input the image photographed by the camera mechanism 29 from the external network and display the image on the display unit 26 through the communication circuit 27 and the signal processing circuit etc.

Beginning at page 50, line 10:

A 17

It is also possible with this HA 1 to photograph a still image by pressing a shutter as in a digital camera for a still image. It is also possible for the camera mechanism 29 to generate a moving picture to output the so-generated moving picture to the signal processor 22. As a signal system for photographing a moving picture by this camera mechanism 29, an MPEG (Moving Pictures Experts Group) system may be used. The camera mechanism 29 provided in the HA 1 is able to photograph a 3D-image to photograph the speaker or the speaker's lip to demonstrate it on the display unit 26 to improve the user's recognition further.

A 18

Beginning at page 51, line 4:

The HA 1 may also be provided with e.g., a switch mechanism so that the user is free to control whether the speech detected by the microphone 21 is outputted to the speech enhancement unit 25, an image photographed by the camera mechanism 29 is to be outputted to the display unit 26 or both the speech and the image are to be outputted. At this time, the switch mechanism, acted on by the user, controls the output from the speech information generating unit 23.

A 19

Beginning at page 53, line 2:

In the foregoing explanation of the HA 1, an exemplary processing of outputting the speech information as the speech is explained. However, this is merely illustrative. For example, the HA 1 may be of the type in which the result of recognition is presented to the user by the middle ear implant. That is, in this HA 1, the speech information may be presented to the user as electrical signals through a coil or a vibrator.

A 20

Beginning at page 53, line 17:

It is also possible for this HA 1 to modulate, work on or transform the result of recognition and the result of recognition worked or transformed into the speech in the